MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2018/2019

BTS3014 –TIME SERIES ANALYSIS AND FORECASTING

(All Sections/ Groups)

6 MARCH 2019
9.00 a.m. - 11.00 a.m.
(2 Hours)

INSTRUCTIONS TO STUDENTS

1. This Question paper consists of **TEN** pages including statistical tables and formulas with **TWO** sections only.
   Section A – Twenty (25) Multiple Choice Questions
   Section B – Three (3) Structured Questions

2. Attempt **ALL** questions in both **Sections A and B**. The distribution of the marks for each question is given.

3. Answer **Section A** in the MCQ answer sheet provided and **Section B** in the answer booklet provided.

4. Statistical tables and formulas are attached.

SECTION A: MULTIPLE CHOICE QUESTIONS (25 MARKS)

1. Of the following model selection criteria, which is often the most important in determining the appropriate forecast method?
   A) Technical background of the forecast user.
   B) Patterns the data have exhibited in the past.
   C) How much money is in the forecast budget?
   D) What is the forecast horizon?

2. Stationarity refers to
   A) the size of the RMSE of a forecasting model.
   B) the size of variances of the model's estimates.
   C) a method of forecast optimisation.
   D) lack of trend in a given time series.

3. The simple exponential smoothing model can be expressed as
   A) a simple average of past values of the data.
   B) an expression combining the most recent forecast and actual data value.
   C) a weighted average, where the weights sum to zero.
   D) a weighted average, where the weights sum to the sample size.

4. How many parameters must the forecaster set using Winter's exponential smoothing?
   A) 0.
   B) 1.
   C) 2.
   D) 3.

5. Which of the following is a tool used in model selection?
   A) Seasonality.
   B) Cyclicity.
   C) Growth.
   D) Plotting the data.

6. First-differencing the data is a way to:
   A) Detrending the data.
   B) Reseasonalising the data.
   C) Removing heteroscedasticity from the data.
   D) Removing any data nonlinearities.

7. Which forecasting model is most likely to appear as a "black box" to management?
   A) Time series decomposition.
   B) Winter's smoothing.
   C) Causal regression.
   D) ARIMA models.

Continued...

8. Which of the following statements are **TRUE**?
   A) Autocorrelation arises when there is a perfect linear association between the dependent and independent variables.
   B) Autocorrelation implies the error terms have differing variances.
   C) Autocorrelation can be tested using the F-statistic.
   D) Autocorrelation causes the estimated regression standard error to be biased.

9. The inclusion of seasonal dummy variables to a multiple regression model may help eliminate
   A) autocorrelation if the data are characterised by seasonal fluctuations.
   B) perfect multicollinearity.
   C) near multicollinearity.
   D) bias in OLS slope estimates caused by autocorrelation.

10. Which of the following is **NOT** recommended to help select the correct set of independent variables for multiple regression?
    A) R-squared.
    B) Adjusted R-squared.
    C) Akaike Information Criterion.
    D) The Bayesian Information Criterion.

11. Including both male and female dummy variables in the same regression will likely result in
    A) near multicollinearity.
    B) perfect multicollinearity.
    C) serial correlation.
    D) heteroscedasticity.

12. In a regression of sales on income and seasonal dummy variables for a quarterly time series, a negative sign of the quarter 3 dummy variable means
    A) sales for quarter three are negative.
    B) sales for quarter three are below average.
    C) sales for quarter three are average below that of the base quarter.
    D) sales for quarter three are above average.

13. Departures from stationarity
    A) jeopardise forecasts and inference based on time series regression.
    B) occur often in cross-sectional data.
    C) can be made to have less severe consequences by using log-log specifications.
    D) cannot be fixed

Continued...

14. One reason for computing the logarithms (ln), or changes in logarithms, of economic time series is that
    A) numbers often get very large.
    B) economic variables are hardly ever negative.
    C) they often exhibit growth that is approximately exponential.
    D) natural logarithms are easier to work with than base 10 logarithms.

15. An autoregression is a regression
    A) of a dependent variable on lags of regressors.
    B) that allows for the errors to be correlated.
    C) model that relates a time series variable to its past values.
    D) to predict sales in a certain industry.

16. The forecast is
    A) made for some date beyond the data set used to estimate the regression.
    B) another word for the OLS predicted value.
    C) equal to the residual plus the OLS predicted value.
    D) close to 1.96 times the standard deviation of Y during the sample.

17. A seasonal index number of 0.80 for quarter one of an automobile parts manufacturer suggests
    A) Quarter one sales are 80% above the norm.
    B) Quarter one sales are 1.80% below the norm.
    C) Quarter one sales are 20% below the norm.
    D) Quarter one sales are 80% below the norm.

18. In computing a seasonal index, specific seasonals were tabulated for each month. The averages over time for the twelve months were obtained and summed. If the mean seasonal factor for June was 96.9, and the sum for all twelve months is 1195; the adjusted seasonal index for June is:
    A) 102.7.
    B) 96.4.
    C) 96.9.
    D) 97.7.

19. What problem will arise when applying ARIMA-type models to highly seasonal monthly data?
    A) Autocorrelation.
    B) Heteroscedasticity.
    C) Stationarity.
    D) Extremely high-order AR and MA processes.

Continued...

20. Which of the following rules is **NOT** a useful first step in the ARIMA model selection process?
    A) If the autocorrelation function stops after q spikes, the appropriate model is a MA(q) type.
    B) If the partial autocorrelation function stops after p spikes, then the appropriate model is an AR(p) type.
    C) If the autocorrelation function does not rapidly approach zero, then first-difference the data.
    D) If the partial autocorrelation function quickly approaches zero, then first difference the data may be recommended.

21. The constant-rate-of-growth model
    A) cannot be estimated with ordinary least squares.
    B) is specified as Y = a + b(TIME).
    C) is specified as $Y = ab^{TIME}$.
    D) requires the trend in a time series to be linear.

22. Forecast errors
    A) are inevitable, so discard them.
    B) are inevitable, so they are not worth discussing with management.
    C) are useful to the extent that they contain valuable information.
    D) will get you fired!

23. Which of the following models utilises a transformed series to induce a stationary series?
    A) ARIMA(1, 0, 1).
    B) ARIMA(1, 0, 0).
    C) ARIMA(1, 1, 1).
    D) ARIMA(0, 0, 1).

24. In preparing the actual forecast numbers, the forecast staff should
    A) prepare a range of possible forecasts.
    B) prepare forecasts using different methods.
    C) examine combined forecasting methods.
    D) All of the above.

25. Stock price data show periods of relatively calm interrupted by periods of enhanced price volatility. This suggests stock price data are
    A) homoscedastic.
    B) autocorrelated.
    C) nonlinear.
    D) heteroscedastic.

Continued...

SECTION B: STRUCTURED QUESTIONS (75 MARKS)

Question One (25 marks)

(a) Define each of the components of the classic time-series multiplicative decomposition method and explain how they are determined. (12 marks)

(b) Describe the family of autoregressive (AR) models and explain the reason why they might be appropriate for the modeling of economic time series. In your answer, discuss the identification and estimation of such models. (7 marks)

(c) Why is it important to test for non-stationarity in time series data before attempting to build an empirical model? Explain. (6 marks)

Question Two (25 marks)

(a) Suppose the following is the monthly total sales data for KFG Company in year 2018.

Month	Sales (in RM thousands)
Jan	6
Feb	10
Mar	10
Apr	8
May	8
Jun	12
July	10
Aug	10
Sep	12
Oct	12
Nov	14
Dec	16

(i) Use four-term moving average to generate one-step-ahead forecasts for May until December 2018. (4 marks)

(ii) Based on the data above, forecast the KFG Company's total sales from May to December 2018 using the simple exponential smoothing method (assume $\alpha = 0.6$, and initial value = 8). (4 marks)

(iii) Evaluate these forecasting methods (from parts (i) to (ii)) using Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). Show all calculations. (6 marks)

Continued...

(iv) Based on part (iii), suggest and explain which forecasting method performs better. Forecast the sales for January 2019 using your suggested method.

(3 marks)

(b) In a time-series multiplicative decomposition of sales (in millions of units), the centered moving-average trend (CMAT) has been estimated:

CMAT = 4.7 + 0.37(T)

The seasonal indices have been found to be:

Quarter	Seasonal Index
1	1.24
2	1.01
3	0.76
4	0.99

For the coming year the time index (T) and cycle factors (CF) are:

Quarter	T	CF
1	21	1.01
2	22	1.04
3	23	1.06
4	24	1.04

From this information prepare a forecast for each quarter of the coming year.

(8 marks)

Question Three (25 marks)

(a) Based on a time-series of quarterly sales of ABC (2009Q1-2018Q4), you estimate the least-square trend line in the form:

SALES = a + b(TIME).

The estimation results are reported below:

Coefficient Table (Linear Regression Selected)

Name	Coefficient Value	Standard Error	t-test	Elasticity	F-test
Intercept	88,741.01	8,043.91	11.03		246.00
Slope	5,362.62	341.91	15.68	0.55	

Continued...

Statistics	
Mean Absolute Percentage Error (MAPE)	12.24%
R-square	86.62%
Durbin Watson	1.55

(i) Write the fitted regression equation. (3 marks)

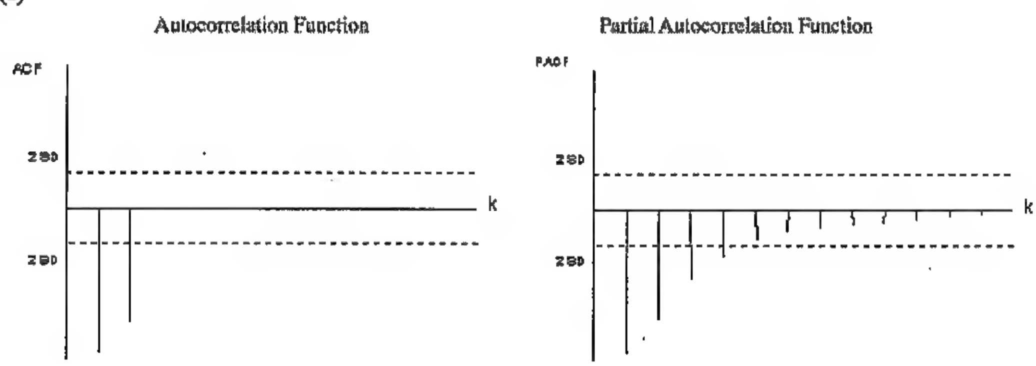
(ii) Do the regression results indicate to you that there is a significant trend to the data? Explain why or why not. (6 marks)

(iii) On the basis of your results, prepare a forecast of sales for the four quarters of 2019. (4 marks)

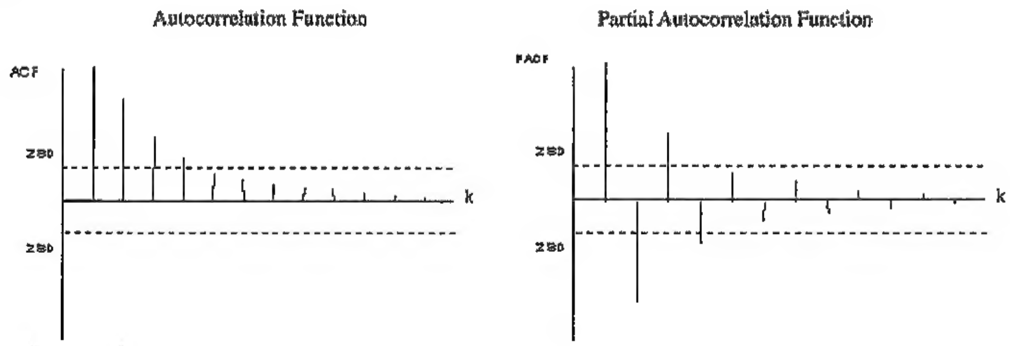
(b) Are ARMA(p, q) models appropriately applied to data with strong trend and pronounced seasonality? Explain how you would apply the Box-Jenkins methodology to such a time series. (8 marks)

(c) Given the graphs of Figure (i)-(ii) of the sample autocorrelations and the sample partial autocorrelations, tentatively identify an ARMA model from each pair of graphs. (4 marks)

(i)



(ii)



End of Page

Statistical Tables

Appendix 1: t-Table

two tails	0.2	0.1	0.05	0.02	0.01
One tail	0.1	0.05	0.025	0.01	0.005
df					
10	1.37	1.81	2.23	2.76	3.17
20	1.33	1.72	2.09	2.53	2.84
30	1.31	1.70	2.04	2.46	2.75
40	1.30	1.68	2.02	2.42	2.70
50	1.30	1.68	2.01	2.40	2.68
60	1.30	1.67	2.00	2.39	2.66
75	1.29	1.67	1.99	2.38	2.64
100	1.29	1.66	1.98	2.36	2.63
120	1.29	1.66	1.98	2.36	2.62
200	1.29	1.66	1.98	2.36	2.62

Appendix 2: Durbin Watson d Table

Critical Values for the Durbin-Watson Statistic (d)

Level of Significance $\alpha = 0.05$

n	$k=1$		$k=2$		$k=3$		$k=4$		$k=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
10	0.88	1.32	0.70	1.64	0.53	2.02	0.38	2.41	0.24	2.82
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
25	1.29	1.41	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78
150	1.72	1.75	1.71	1.76	1.69	1.77	1.68	1.79	1.66	1.80
200	1.76	1.78	1.75	1.79	1.74	1.80	1.73	1.81	1.72	1.82

where n = number of observations and k = number of independent variables, excluding the intercept.

Formulas

Error, $\varepsilon_t = y_t - \hat{y}_t$

Mean Absolute Error, $\text{MAE} = \frac{1}{n}\sum_{t=1}^{n}|\varepsilon_t|$

Mean Squared Error, $\text{MSE} = \frac{1}{n}\sum_{t=1}^{n}\varepsilon_t^{\,2}$

Root Mean Squared Error, $\text{RMSE} = \sqrt{MSE}$

Mean Absolute Percentage Error,

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}|PE_t| = \frac{1}{n}\sum_{t=1}^{n}\left|\left(\frac{\varepsilon_t}{y_t}\right)x100\right|$$

Mean Percentage Error,

$$\text{MPE} = \frac{1}{n}\sum_{t=1}^{n}PE_t = \frac{1}{n}\sum_{t=1}^{n}\left(\frac{\varepsilon_t}{y_t}\right)x100$$

Theil's U-statistics, $U = \sqrt{\dfrac{\sum_{t=1}^{n}\varepsilon_t^2}{\sum_{t=1}^{n}(y_t - y_{t-1})^2}}$

Naïve model, $\hat{Y}_{t+1} = Y_t$

Naïve trend model,
$\hat{Y}_{t+1} = Y_t + (Y_t - Y_{t-1})$

Naïve rate of change model,
$\hat{Y}_{t+1} = Y_t * \frac{Y_t}{Y_{t-1}}$

Naïve seasonal model for quarterly data,
$\hat{Y}_{t+1} = Y_{t-3}$

Naïve trend and seasonal model for quarterly data, $\hat{Y}_{t+1} = Y_{t-3} + \frac{(Y_t - Y_{t-4})}{4}$

Simple average model, $\hat{Y}_{t+1} = \frac{1}{t}\sum_{i=1}^{t}Y_t$

Moving average model for k time periods,
$\hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + ... + Y_{t-k+1}}{k}$

Double moving average,

$$M_t = \hat{Y}_{t+1} = \frac{Y_t + Y_{t-1} + ... + Y_{t-k+1}}{k}$$

$$M_t' = \frac{M_t + M_{t-1} + M_{t-2} + ... + M_{t-k+1}}{k}$$

$$a_t = M_t + (M_t - M_t') = 2M_t - M_t'$$

$$b_t = \frac{2}{k-1}(M_t - M_t')$$

$$\hat{Y}_{t+p} = a_t + b_t p$$

Simple exponential smoothing,
$\hat{Y}_{t+1} = \alpha Y_t + (1-\alpha)\hat{Y}_t$

Holt's linear smoothing

The exponentially smoothed series, or current level estimate,
$L_t = \alpha Y_t + (1-\alpha)(L_{t-1} + T_{t-1})$

The trend estimate,
$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$

Forecast p periods into the future,
$\hat{Y}_{t+p} = L_t + pT_t$

Holt-Winter's multiplicative smoothing

The exponentially smoothed series or level estimate,

$$L_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + T_{t-1})$$

The trend estimate,

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1}$$

The seasonality estimate,

$$S_t = \gamma \frac{y_t}{L_t} + (1-\gamma)S_{t-s}$$

Forecast p periods into the future,

$$\hat{Y}_{t+p} = (L_t + pT_t)S_{t-s+p}$$

Time additive decomposition,

$$Y_t = S_t + T_t + C_t + I_t$$

Time series multiplicative decomposition,

$$Y_t = S_t * T_t * C_t * I_t$$

AR process: p-order autoregressive model

$$y_t = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 y_{t-2} + \dots + \gamma_p y_{t-p} + \varepsilon_t$$

MA process: q-order moving average model

$$y_t = \mu_t + \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \dots - \theta_q\varepsilon_{t-q}$$

Regression model:

$$\text{t - statistic} = \frac{\hat{\beta} - \beta}{se(\hat{\beta})}$$